DOCUMENT RESUME

ED 264 441 CE 043 178

AUTHOR Askov, Eunice N.; And Others
TITLE Single-Subject, Multiple-Baseline Designs in Evaluating Adult Literacy Programs.
PUB DATE Nov 85
NOTE 9p.; Paper presented at the National Adult Education Conference (Milwaukee, WI, November 1985).
PUB TYPE Viewpoints (120) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Adult Literacy; *Adult Programs; Educational Assessment; *Educational Research; *Evaluation Methods; Postsecondary Education; *Research Design; *Research Methodology; Research Needs
IDENTIFIERS *Single Subject Research Design

ABSTRACT

Between-group designs constitute the experimental methods traditionally applied in the field of education. Another level of analysis, however, is commonly used in clinical psychology but seldom applied in educational settings: single-subject designs (actually applied to small groups). This design allows a researcher to evaluate the influence of an educational or clinical procedure on a subject's performance by replicating its effect with the same individual over time. One type of single-subject design, the multiple-baseline design, is becoming more popular in educational studies. In the multiple-baseline design, measures are made simultaneously on two or more behaviors, subjects, or settings in order to identify a pre-treatment (baseline) level. The baseline data describe the current level of performance and predict future performance. If each baseline changes when the intervention is introducted, the effects can be attributed to the intervention. Advantages of multiple-baseline designs include the following: (1) the multiple-baseline designs are effective means of evaluating educational procedures since they allow continuous monitoring (2) intervention is applied to only one or two baselines at a time; (3) large numbers of subjects are not needed; (4) the gradual application of the intervention permits the researcher to test its effectiveness on a small scale and make needed modifications; and (5) establishing a causal relationship does not require the withdrawal of treatment. Multiple-baseline designs may be appropriate for researching a variety of educational questions, such as evaluation of adult literacy programs. (KC)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

SINGLE-SUBJECT, MULTIPLE-BASELINE DESIGNS IN EVALUATING ADULT LITERACY PROGRAMS

Eunice N. Askov
Sarah D. Weidler
Connie M. Maclay
The Pennsylvania State University
Institute for the Study of Adult Literacy

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
E Askov
S Weidler, C Maclay
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

ED264441

Between-group designs constitute the experimental methods traditionally applied in the field of education. Frequently one method of instruction is compared to another, or different conditions of instruction are compared, for two groups of subjects. Treatment is sometimes withheld for a control group to permit comparison to an experimental group receiving the treatment. Withholding treatment, such as keeping adult literacy students on a class waiting list, creates ethical problems as well as recruitment difficulties.

Another level of analysis, however, is commonly used in clinical psychology but seldom applied in educational settings: single-subject designs. This method of research presents an alternative approach to between-group research. Although there are a number of single-subject designs available, for the purpose of this paper only single-subject, multiple-baseline designs are presented.

The term, single-subject designs, is really an inaccurate label since most researchers apply the design to small groups of subjects. The designs are intended to identify and maximize the uniqueness of the individuals within the groups (Schloss, Sedlak, Elliot, and Smothers, 1982). The designs allow a researcher to evaluate the influence of an educational or clinical procedure on a subject's performance by replicating its effect with the same individual over time. Essentially, the subjects function as their own controls.

One type of single-subject design, the multiple baseline design, is becoming more popular in educational studies. In the multiple-baseline design, measures are made simultaneously on two or more behaviors, subjects, or settings in order to identify a pre-treatment (baseline) level. The baseline data describes the current level of

CE043178

performance and predicts future performance. If each baseline changes when the intervention is introduced, the effects can be attributed to the intervention. Causality is established by the change in one baseline (when the intervention is introduced) while there is no concurrent change in another baseline (when the intervention has not been introduced) (Kazdin, 1982).

Replication of the results is achieved by applying the treatment to the additional baseline(s). This repeated demonstration that behavior changes in response to the intervention usually makes implausible the influence of extraneous factors (Kazdin, 1982; Tawney & Gast, 1984). The different baselines in the design serve as control conditions to evaluate what changes can be expected without the treatment.

One major advantage of the multiple-baseline design over between-group designs is its flexibility. Individual attention can be given to each client with opportunities to alter and improve the intervention when needed. This is particularly acceptable if the intervention has not been applied in a specific research situation or if it is new or untried (Kazdin, 1982).

There are three common variations of the multiple-baseline design: multiple-baseline across individuals, behaviors, and conditions. In the multiple-baseline across individuals, baseline data are gathered for one specific behavior which is performed by two or more individuals. The baselines refer to the separate individuals within the study. In most designs across subjects, the researcher identifies two or more subjects who demonstrate the same behavior. Baseline observation data are collected on these indivduals. The intervention is applied to each subject in turn while data are continuously collected on all subjects. The effect of the intervention is demonstrated when the change in each baseline coincides with the point when the intervention was introduced and not before (Halle, Stoker, and Schloss, 1984; Kazdin, 1982; Tawney and Gast, 1984).

In the multiple-baseline across behaviors, the researcher identifies two or more independent yet similar observable behaviors of an individual or group of individuals. Each target behavior is measured simultaneously and continuously under the same conditions until a sufficiently stable baseline is identifed. The intervention (independent variable) is sequentially applied to the target behaviors. This design is especially effective in evaluating an intervention intended to either increase the frequency of an approriate academic or social behavior or decrease the frequency of an inappropriate academic or social behavior (Tawney and Gast, 1984).

In the multiple-baseline design across conditions, baseline data are gathered for one specific behavior which is performed by one individual or group (which is considered a unit). The multiple baselines refer to "the different situations, settings, or time periods of the day in which observations are obtained" (Kazdin, 1982, p. 134). The design involves observation of the target behavior in each situation. After a baseline has been established for each situation, the intervention is applied sequentially while continuous measuring of the baseline is carried out across all baselines.

Table 1 may be helpful in summarizing the three types of multiple-baseline designs. Though these three represent the most common currently found in the literature, combinations and variations are also available.

There are several prerequisite steps which may prove useful when implementing a multiple-baseline design (Kazdin, 1982; Halle, Stoker, and Schloss, 1984): (1) the target behavior should be defined; (2) an appropriate recording system should be selected; and (3) the reliability of the measures should be computed. In order to operationally define the target behavior, it must be both observable and measurable. Its definition should meet several criteria: objectivity, clarity, and completeness (Kazdin, 1982). Halle, Stoker, and Schloss (1984) identify commonly used assessment strategies, such as frequency counts (e.g., the number of words identified correctly) and latency (e.g., the time needed to identify a word).

Of prime importance is the computation of some sort of interobserver agreement in order to confirm or disconfirm the reliability of the assessment. If frequency counts vary from one observer to another, it would be difficult to ascertain the subject's actual performance. A reliability check between observers also helps to minimize observer biases and lends support for future replication by showing that the target behavior definitions are complete and accurate. For a more comprehensive explanation of defining target behavior, selecting appropriate behavioral assessment, and calculating interobserver agreement, refer to Kazdin (1982).

Typically, visual inspection is the manner in which any intervention effects are noted for single-subject research. The effects of the intervention are considered clear if there is a noticable shift in the level of the graphed data points when the intervention is applied. The data are plotted graphically so that patterns which might indicate a causal relationship can be inferred (Kazdin, 1982). The more closely in time that the hypothesized change occurs after the experimental conditions have been altered, the clearer the intervention effects. As a general rule, the shorter the period of time between the introduction of the intervention and the behavior change, the easier it is to infer that the intervention led to change; it is less likely that any other intervening influences have accounted for the change.

Some limitations exist in the multiple-baseline design. In some studies it is quite difficult to identify similar yet independent behaviors or conditions. Interdependence of baselines can confound the results of the treatment (Kazdin, 1982). This can especially be a problem in multiple-baseline designs across subjects since the change in one individual's behavior may influence others who have not yet received the intervention (e.g., in an adult basic education class setting). In such a situation, the relationship of the intervention and the behavior change is ambiguous.

Another limitation is the possibility of inconsistent effects of the intervention (Kazdin, 1982). Some behaviors may change when the intervention is introduced while

others may not. This raises questions about the generality or strength of the intervention. One should keep in mind that not all baselines need to show the predicted change in order to infer that the intervention was responsible for the change. If several of the baselines show the predicted effect, an exception may not negate the drawing of inferences but may, instead, suggest areas for further study.

A third limitation is the possibility of a prolonged baseline which could lead to methodological problems in that the behaviors may sometimes improve even before the intervention is introduced. Prolonging baselines may also be time-consuming, impractical, distracting, and costly (Tawney and Gast, 1984). As with the no-treatment control group, the ethical question of withholding treatment exists. Some variations have been proposed to avoid this situation, such as introducing the treatment to two behaviors at the same time or observing at longer intervals rather than daily or hourly.

A fourth limitation is the possibility of a treatment x subject interaction. In other words, some unique characteristics of the subject may have made the treatment more or less effective. If that characteristic can be identified in the subject, and in other potential students, then the treatment can be used (or not used) with those having that characteristic. For example, a particular type of literacy instruction may not be effective with learning disabled adult students. For those students, that type of treatment would not be used.

The advantages of the multiple-baseline designs are summarized below:

1. The multiple-baseline designs are effective means of evaluating educational procedures since they allow continuous monitoring through both learning and retention stages (Tawney and Gast, 1984).

2. The intervention is applied to only one (or two) baselines at a time allowing the researcher to focus more closely on the changes which may occur (Halle, Stoker, and Schloss, 1984).

3. The multiple-baseline designs are not dependent upon large numbers of subjects.

4. The gradual application of the intervention permits the researcher to test its effectiveness on a small scale and make modifications if necessary (Kazdin, 1982).

5. The multiple-baseline designs do not require the withdrawal of treatment to establish a causal relationship.

The multiple-baseline designs may be appropriate for researching a variety of educational questions. They allow the researcher to look closely at an intervention and the applicability to a variety of behaviors, conditions, and subjects -- clearly a useful tool which is well suited for both clinical and classroom research. These designs can be very useful in evaluating adult literacy programs when use of a control group is not possible. They can also be used in evaluating pilot studies which precede a large scale, between-group study.

References

Halle, J. W., Stoker, R., and Schloss, P. J. (1984). Facilitating teacher-conducted research: A tutorial on single-subject design - the multiple-baseline. The Volta Review, 86, (2), 89-101.

Kazdin, A. E. (1982). Single-case research designs: Methods for clinical and applied settings. New York: Oxford University Press.

Schloss, P. J., Sedlak, R. A., Elliot, C., and Smothers, M. (1982). Application of the changing-criterion design through token reinforcement. Journal of Special Education, 16, (3), 359-367.

Tawney, J. W., and Gast, D. (1984). Single-subject research in special education. Columbus, Ohio: Charles E. Merrill.

Table 1

Multiple Baseline Designs

	The researcher must identify two or more different...	The researcher will observe the same...	This design helps to measure...
Across behaviors	behaviors	condition(s) and subject(s)	the effectiveness of the intervention in changing more than one similar yet independent behaviors.
Across conditions	conditions	behavior(s) and subject(s)	the effectiveness of the intervention in changing the target behavior in a variety of situations, settings, and times.
Across subjects	subjects	behavior(s) and condition(s)	the effectiveness of the intervention in changing the target behavior in two or more subjects.